United Nations E/CN.3/2011/30

Economic and Social Council

Distr.: General
8 December 2010

Original: English

Statistical Commission
Forty-second session
22-25 February 2011
Item 4 (j) of the provisional agenda*
Items for information: inventory of global statistical standards

Report of the Committee for the Coordination of Statistical Activities on the inventory of global statistical standards

Note by the Secretary-General

In accordance with a request of the Statistical Commission at its forty-first session (see E/2010/24, chap. I.B), the Secretary-General has the honour to transmit the report of the Committee for the Coordination of Statistical Activities on its progress in establishing an inventory of global statistical standards, based on a system originally developed by the National Institute of Statistics and Geography of Mexico. The Commission is invited to take note of this progress report.

* E/CN.3/2011/1.

10-68225 (E) 201210

Please recycle

Report of the Committee for the Coordination of Statistical Activities on the inventory of global statistical standards

I. Introduction

1. At its forty-first session the Commission considered a report prepared by the National Institute of Statistics and Geography of Mexico (INEGI) on its efforts to elaborate an inventory of international statistical standards (E/CN.3/2010/14). The development of this inventory had been motivated by a requirement of the new Mexican statistical law that national methodologies concord with international recommendations and standards. Mexico offered this system to the international community as a useful frame of reference. In addition, it was suggested that there be a process for reviewing, validating and updating the system in the future, in close collaboration with the international organizations that are responsible for maintaining the various standards in the different subject matter areas.

2. The Statistical Commission, in its decision 41/111, welcomed the proposal of the Committee for the Coordination of Statistical Activities to create appropriate mechanisms for the further development and maintenance of the inventory of global statistical standards and requested it to propose a proper classification of those standards, and also requested it to submit a progress report to the Commission at its forty-second session.

II. Review of the system by the Committee for the Coordination of Statistical Activities

3. The Committee for the Coordination of Statistical Activities first considered the topic of a global inventory of statistical standards at its fifteenth session, held in New York on 22 February 2010 (see E/CN.3/2011/29), on the basis of a paper and a presentation by a representative of INEGI. The Committee noted that the further development and maintenance of the inventory would require significant efforts on the part of all its members. In the absence of a definition of what constituted an "international statistical standard", the Committee decided that it was imperative to look into that issue and devise a practical proposal regarding the coverage of such a global database. Questions were also raised regarding the inclusion of "regional standards", the translation of the entire database and the need for a classification of statistical domains. The Committee requested a group composed of representatives of the Economic Commission for Europe (ECE), the United Nations Statistics Division, the World Bank and the International Labour Organization (ILO), in collaboration with INEGI, to prepare and present to its sixteenth session a concept paper on how to develop the global inventory further, including proposals on how to ensure its future maintenance.

4. At its sixteenth session, held in Vienna from 1 to 3 September 2010 (ibid.), the Committee considered this issue further on the basis of a report prepared by ECE on behalf of the task team that had analysed the following aspects of the system and presented its conclusions thereon.

A. Purpose

5. The inventory is intended to provide a knowledge centre and reference system for international statistical standards. It therefore needs to be complete, up-to-date, trustworthy and readily available for users.

6. As a reference system, the inventory can also serve as an analytical tool for international organizations and groups, to help identify statistical areas requiring recommendations or standards as well as to identify overlapping standards. It will therefore provide the basis for the Statistical Commission to manage the system of global statistical standards.

7. It is proposed that the inventory be developed in phases. The first phase will provide a system that meets the main requirements, with a limited level of detail. User feedback will help to determine the content and priorities of future phases, though ideas for future work are included in some of the following sections.

B. Users

8. As well as the Statistical Commission and international organizations, the inventory is expected to be used by professionals in national statistical organizations, the research community and the general public.

C. Coverage

9. The main issues concerning coverage are, firstly, to define what is meant by a "statistical standard" and secondly, to determine the scope of the standards to be included in the inventory. Concerning the definition of a statistical standard, it is proposed to follow the one given in annex 4 — Metadata Common Vocabulary to the Statistical Data and Metadata eXchange (SDMX) Content-oriented Guidelines, 2009[1]: "An agreed rule or guideline on how one or more parts of the statistical business process should be carried out, conforming with requirements for professionalism".

10. In this context, "an agreed rule or guideline" should be interpreted as one which has been adopted by an international body and/or published by an international organization. The "requirements for professionalism" are those enshrined in the Fundamental Principles of Official Statistics, the Principles Governing International Statistical Activities and the International Statistical Institute Declaration on Professional Ethics.

11. Concerning the scope of standards, it is proposed to include:

- All official methodological publications of Committee members
- Internationally agreed statistical classifications
- Standards published by a body outside the international statistical system, but widely used by statistical organizations (e.g. some International Organization for Standardization standards)

[1] Available from http://sdmx.org/?page_id=11.

- Global and regional standards adopted by the relevant global or regional statistical bodies, including regional versions of global standards
- Information about previous versions of standards where these are still used

12. Standards that are used only within one country or organization will not be included in the inventory.

13. In the first phase of the project, the main entry for the inventory should be the official publication rather than the individual concepts and definitions included therein. However, this may not provide sufficient detail for the purposes of checking the coherence of standards and measuring compliance, so it may be necessary to move to a more detailed level later.

D. Classification

14. It is proposed to use the classification of statistical activities developed for the ECE Database of International Statistical Activities (DISA). This list of statistical activities is managed by the Bureau of the Conference of European Statisticians and has also been adopted as a standard by other organizations and initiatives (e.g. SDMX and PARIS21).

15. It is also proposed to relax the usual rule that items can only be allocated to one category of a classification, in the interests of helping users find relevant standards. For example, the "System of National Accounts" is most relevant to category 2.2 (Economic accounts) of the classification of statistical activities, but is also relevant to several other categories. A reference system based on keywords related to one or more categories could be envisaged in future phases of the development of the inventory.

E. Information contents

16. The inventory should contain structured entries describing each standard and providing links to further information. After having reviewed the metadata model used by INEGI for its inventory and the template used in the Conference of European Statisticians' Common Metadata Framework inventory of standards related to statistical metadata,[2] the information content model as shown in the following table was proposed for the inventory.

[2] The Common Metadata Framework concentrates on different practical and theoretical aspects of statistical metadata systems and provides vital knowledge for anyone working with statistical metadata. Part B deals with metadata concepts, standards, models and registries, and can be found at www.unece.org/stats/cmf.

Table
Proposed information content model for the Global Inventory of Statistical Standards

Element	*Description*
Name	Full official name of the standard, plus any alternative names and acronyms
Version	The version of the standard (where applicable)
Valid from	Date on which the standard was adopted (or published)
Previous versions	References and links to previous versions of the standard
Description	Brief description of the standard, including objective, intended use and relationship to other standards
Maintenance organization	Organization or group responsible for approving and maintaining the standard
Statistical domain(s)	Domains in which the standard is used (according to the DISA classification of statistical activities)
Type of standard	Categories: (a) Concepts and definitions; (b) Classifications; (c) Methodologies and procedures. (*Note*: other categories could be added in the future if needed.)
References	References and links to the standard and any other material that helps understanding or implementation
Language	Language(s) in which the standard is published
Information last updated/validated	Date on which the above information was last updated or verified to be correct

F. Functionality

17. The three main use-cases envisaged at this stage are:

- A user wants to find information about a specific standard
- A user wants to find information about standards relevant for a specific statistical domain or part of the statistical production process
- International organizations want to manage the system of statistical standards (identify gaps and overlaps, measure compliance, etc.)

18. A simple, user-friendly interface is therefore required, allowing standards to be found by using a search function and by using the domain classification. On finding

the description of the standard, the user should be able to easily navigate to the standard itself (if it exists in electronic form on the Internet). There should also be links to supporting material concerning the standard and its implementation, and to related standards.

19. Conceptual maps describing relations between groups of standards could be a useful addition, but are probably not essential for the first phase of the inventory. Work on this sort of mapping is currently under way for metadata-related standards.

20. Functionality for maintenance of the inventory should also be considered. This task is likely to be spread among a number of editors from various international organizations and other authorized bodies in different locations. This will require a remote-access facility with authentication for users granted update rights. Those updating the inventory are unlikely to be information technology specialists, so a user-friendly text-editor interface will be required.

21. The processes and tools for maintaining the inventory are discussed further below.

G. Languages

22. There is a clear need to find a balance between the user requirement for information in many different languages and the substantial costs of managing and ensuring the consistency of a multilingual system. While it is relatively easy to manage multiple-language versions of the headings in a template, it is less easy to manage free-text fields. It might therefore be preferable to structure information in the template as much as possible (for example, using options in drop-down menus rather than free-text fields). The feasibility of linking free-text fields to an automatic translation engine (e.g. Google Translate or Babelfish) should also be considered. While automatic translations are often far from perfect, they are gradually improving and may be sufficient in this case.

H. Maintenance

23. The resource requirement for maintaining an inventory of this type should not be underestimated. A long-term commitment to providing sustainable resources should be made before the project starts, to avoid the result becoming just another out-of-date website.

24. It is also likely that a certain amount of domain-specific knowledge will be needed to maintain the inventory. A single administrator could not be expected to follow developments in all statistical domains. For this reason, it is proposed to appoint a domain editor for each statistical domain. This person will be an expert in his/her field from a national or international statistical organization, and may also represent an international expert group covering the specific domain. One person can cover more than one domain where appropriate.

25. There will also be a requirement for one person (plus back up) to administer the inventory website, including allocating permission to editors and monitoring site usage.

26. It would also help to have an editorial board to take responsibility for matters of content, such as changes to the template, classifying statistical standards by domain and deciding whether specific standards meet the criteria for inclusion in the inventory. This clearly requires an expert-level group with experience of working with statistical standards.

I. Platform

27. The experiences of managing DISA and the Common Metadata Framework inventory of metadata standards suggest that a "wiki" would be the most appropriate technical platform for the inventory. This approach seems to be the simplest, most cost-effective solution. It would allow controlled remote access for editors and provide the other functional requirements identified in the preceding sections.

III. Development of a prototype

28. The Committee recommended that the group working on the inventory should take a pragmatic approach to its development. Specifically, it was proposed that the inventory be developed in phases. The first phase would provide a prototype that met the main requirements and contain a limited level of information detail. User feedback would help determine the priorities for a second or, if necessary, additional future phases. The development of the inventory would be based on the following principles:

- Ownership — while the practical overall execution of the project lies with the United Nations Statistics Division, the inventory will need to be owned in partnership by all CCSA members
- Location — the inventory will be hosted by the United Nations Statistics Division
- Maintenance — substantial investments will be required during the first phase of the project; the follow-up maintenance is expected to be fairly easy. Decentralized data entry with remote update under the coordination of the Division was favoured, with key organizations responsible for their respective subject-matter domains
- Language — the inventory has to be developed first in English, keeping the corresponding existing references in Spanish to the extent possible
- Classification of statistical activities — DISA was considered to be an appropriate starting point for the inventory. Given that the Bureau of the Conference of European Statisticians is the custodian of DISA, all proposals for modifications and/or inclusion of new domains will need to be brought to its attention
- Types of standards — further work is needed to clarify borderline cases of international standards, guidelines, etc.
- Quality of the inventory — this will depend on the metadata references for the documents included

29. The Committee has requested a group composed of ECE, the United Nations Statistics Division, the World Bank, ILO and INEGI to draft more detailed terms of reference describing all the procedures identified in the discussion. The group was also requested to prepare a prototype of the inventory and to present it at the seventeenth session of the Committee to be held in February 2011. ECE and the United Nations Statistics Division have taken the lead in the development of this prototype, which is wiki-based. As there is an increasing trend towards process-based standardization, it is being considered to add a process dimension to the classification of standards, based on the Generic Statistics Business Process Model developed by the ECE/Eurostat/Organization for Economic Cooperation and Development group on statistical metadata.[3] A first version of the prototype will be ready for demonstration by the time of the forty-second session of the Statistical Commission.

[3] See www.unece.org/stats/gsbpm.